

0590

07/17

#6



OIKE

ENTERED

RAW SEQUENCE LISTING

DATE: 07/24/2002

PATENT APPLICATION: US/10/040,315A

TIME: 11:25:15

Input Set : D:\seqlist.txt

Output Set: N:\CRF3\07242002\J040315A.raw

```

4 <110> APPLICANT: Farese, Robert V.
5     Cases, Sylvaine
6     Smith, Steven
7     Erickson, Sandra
10 <120> TITLE OF INVENTION: Diacylglycerol O-Acyltransferase
13 <130> FILE REFERENCE: UCAL-105CIP2
15 <140> CURRENT APPLICATION NUMBER: 10/040,315A
16 <141> CURRENT FILING DATE: 2001-10-29
18 <150> PRIOR APPLICATION NUMBER: 60/107,771
19 <151> PRIOR FILING DATE: 1998-11-09
21 <150> PRIOR APPLICATION NUMBER: PCT/US98/17883
22 <151> PRIOR FILING DATE: 1998-08-28
24 <150> PRIOR APPLICATION NUMBER: 09/103,754
25 <151> PRIOR FILING DATE: 1998-06-24
27 <150> PRIOR APPLICATION NUMBER: 09/339,472
28 <151> PRIOR FILING DATE: 1999-06-23
30 <160> NUMBER OF SEQ ID NOS: 10
32 <170> SOFTWARE: FastSEQ for Windows Version 3.0
34 <210> SEQ ID NO: 1
35 <211> LENGTH: 1411
36 <212> TYPE: DNA
37 <213> ORGANISM: homo sapiens
39 <400> SEQUENCE: 1
40 ttatttttgg agaacctcat caagtatggc atcctggtgg accccatcca ggtggtttct      60
41 ctgttccctga aggatcccta tagctggccc gcccatgcc tggttattgc ggccaatgtt      120
42 tttgctgtgg ctgcattcca ggttgagaag cgctggcgg tgggtgccct gacggagcag      180
43 gcgggactgc tgtgcacgt ggccaacctg gccaccattc tgtgtttccc agcggctgtg      240
44 gtcttaactgg ttgagtctat cactccagtg ggctccctgc tggcgtgat ggcgacacc      300
45 atcctcttcc tcaagctctt ctctaccgc gacgtcaact catggtgccg cagggccagg      360
46 gccaaggctg cctctgcagg gaagaaggcc agcagtgttg ctgccccgca caccgtgagc      420
47 taccgggaca atctgacctc ccgcgatctc tactacttcc tcttcgcccc caccctgtgc      480
48 tacgagctca actttccccg ctctccccgc atccggaagc gctttctgct gcgacggatc      540
49 cttgagatgc tgttcttcac ccagctccag gtggggctga tccagcagtg gatggctccc      600
50 accatccaga actccatgaa gcccttcaag gacatggact actcacgcat catcgagcgc      660
51 ctctgaagc tggcggtccc caatcacctc atctggctca tcttcttcta ctggctcttc      720
52 cactcctgcc tgaatgccgt ggctgagctc atgcagtgtg gagaccggga gttctaccgg      780
53 gactggtgga actccgagtc tgtcacctac ttctggcaga actggaacat ccctgtgcac      840
54 aagtggtgca tcagacatt ctacaagccc atgcttcgac ggggcagcag caagtggatg      900
55 gccaggacag ggggtttcct ggctcggcc ttctccacg agtacctgtg gagcgtccct      960
56 ctgcgaatgt tccgcctctg ggcgttcacg ggcatgatgg ctacagatccc actggcctgg      1020
57 ttcgtgggcc gctttttcca gggcaactat ggcaacgcag ctgtgtggct gtcgctcatc      1080
58 atcgacagc caatagccgt cctcatgtac gtccacgact actacgtgct caactatgag      1140
59 gccccagcgg cagaggcctg agctgcacct gaggggctgg cttctcactg ccacctcaca      1200

```

RAW SEQUENCE LISTING

DATE: 07/24/2002

PATENT APPLICATION: US/10/040,315A

TIME: 11:25:15

Input Set : D:\seqlist.txt

Output Set: N:\CRF3\07242002\J040315A.raw

```

60 cccgctggca gagccacct ctctcctag gcctcgagtt gctggggatg ggcctggctg 1260
61 cacagcatcc tcctctggtc ccaggaggc ctctctgccc ctatggggct ctgtcctgca 1320
62 cccctcaggg atggcgacag caggccagac acagtctgat gccagctggg agtcttgctg 1380
63 accctgcccc gggtcggagg gtgtcaataa a 1411
65 <210> SEQ ID NO: 2
66 <211> LENGTH: 261
67 <212> TYPE: DNA
68 <213> ORGANISM: homo sapiens
70 <400> SEQUENCE: 2
71 cagaggcctg agctgcacct gaggggctgg cttctcactg ccacctcaca cccgctggca 60
72 gagccacct ctctcctag gcctcgagtt gctggggatg ggcctggctg cacagcatcc 120
73 tcctctggtc ccaggaggc ctctctgccc ctatggggct ctgtcctgca cccctcaggg 180
74 atggcgacag caggccagac acagtctgat gccagctggg agtcttgctg accctgcccc 240
75 gggtcggagg gtgtcaataa a 261
77 <210> SEQ ID NO: 3
78 <211> LENGTH: 1650
79 <212> TYPE: DNA
80 <213> ORGANISM: mus musculus
82 <400> SEQUENCE: 3
83 ggatgaatgg aaataagtag aattaggcat acttaggata gggctcaagc cgcggcccg 60
84 gaagattggg ccgcgacgag gtgcgggccc aagccatggg cgaccgcgga ggcgcgggaa 120
85 gctctcggcg tcggaggacc ggctcgcggg ttcccgctca ggggtgtagt gggcccaagg 180
86 tagaagagga cgaggcgca gacgcggctg tgagccccga cttgggcgcc gggggtgacg 240
87 cgccggctcc ggctccggct ccagcccata cccgggacaa agacgggcgg accagcgtgg 300
88 gcgacggcta ctgggatctg aggtgccatc gtctgcaaga ttctttgttc agctcagaca 360
89 gtggtttcag caattatcgt ggtatcctga attggtgtgt ggtgatgctg atcctgagta 420
90 atgcaagggt atttttagag aacctatca agtatggcat cctggtggat cctatccagg 480
91 tgggtgtctc gtttttgaag gaccctaca gctggcctgc cccatgcgtg attattgcat 540
92 ccaatatatt tgttgtggct gcatttcaga ttgagaagcg cctggcagtg ggtgcctga 600
93 cagagcagat ggggctgctg ctacatgttg ttaacctggc cacaatcatt tgcttcccag 660
94 cagctgtggc cttactgggt gactctatca ctccagtggt ttccgtgttt gctctggcat 720
95 catactccat catgttcctc aagctttatt cctaccggga tgtcaacctg tggtgccgcc 780
96 agcgaagggt caaggccaaa gctgtctcta cagggaagaa ggtcagtggt gctgctgcc 840
97 agcaagctgt gagctatcca gacaacctga cctaccgaga tctctattac ttcatctttg 900
98 ctccactttt gtgttatgaa ctcaactttc ctcggtcccc cgcaatacga aagcgctttc 960
99 tgctacgacg agttcttgag atgctctttt ttaccagct tcaagtgggg ctgatccaac 1020
100 agtggatggt ccctactatc cacaactcca tgaagccctt caaggatatg gactattcac 1080
101 ggatcattga gcgtctctta aagctggcgg tccccaacca tctgatctgg cttatcttct 1140
102 tctattgggt ttccactcc tgtctcaatg ctgtggcaga gcttctgcag tttggagacc 1200
103 gcgagttcta cagagattgg tggaatgctg agtctgtcac ctacttttg cagaactgga 1260
104 atatccccgt gcacaagtgg tgcacagac acttctacaa gcctatgctc agacatggca 1320
105 gcagcaaagt ggtggccagg acaggagtat ttttgacctc agccttcttc catgagtacc 1380
106 tagtgagcgt tcccctgcgg atgttccgcc tctgggcatt cacagccatg atggctcagg 1440
107 tccactggc ctggattgtg ggccgattct tccaagggaa ctatggcaat gcagctgtgt 1500
108 gggtgacact catcatgggg caaccgggtg ctgtgctcat gtatgtccac gactactacg 1560
109 tgctcaacta cgatgcccc gtgggggtat gagctactgc caaaggccag ccctccctaa 1620
110 cctgggcctg gagttctgga ggggttccctg 1650
112 <210> SEQ ID NO: 4
113 <211> LENGTH: 629

```

RAW SEQUENCE LISTING

DATE: 07/24/2002

PATENT APPLICATION: US/10/040,315A

TIME: 11:25:15

Input Set : D:\seqlist.txt

Output Set: N:\CRF3\07242002\J040315A.raw

114 <212> TYPE: DNA

115 <213> ORGANISM: arabidopsis thaliana

117 <220> FEATURE:

118 <221> NAME/KEY: misc_feature

119 <222> LOCATION: (0)...(0)

120 <223> OTHER INFORMATION: Each n residue at position 455, 464, 467, 475, 497, 500,
508,

121 514, 519, 536, 543, 544, 576, 583, 584 and 597 can be either a, c, g or t

123 <400> SEQUENCE: 4

124 tgcattgtata cgggaagggtt ggggtggctcg tcaatttgca aaactgggtca tattcaccgg 60

125 attcatggga ttataatag aacaatatat aaatcctatt gtcaggaact caaagcatcc 120

126 ttgaaaggc gatcttctat atgctattga aagagtgttg aagctttcag ttccaaattt 180

127 atatgtgtgg ctctgcatgt tctactgctt cttccacctt tggtaaaaca tattggcaga 240

128 gcttctctgc ttcggggatc gtgaattcta caaagattgg tggaatgcaa aaagtgtggg 300

129 agattactgg gagaatgtgg aatatgcctg tccataaatg ggatgggtcc gacatatata 360

130 ccttccccgt gcttgccgac aaggattacc caaagacacc cgggccatta accattggct 420

W--> 131 ttcccaagcc ccttggaggc ctttccatgg gccanggacc cggngtncce tggcngggccc 480

W--> 132 ttcaaagcaa agggggnttn cctggggnta aagntccang ggcccttggg gcccanccaa 540

W--> 133 aannttcccc cgggaaaggg ttgccaccg gggggngaaa aanncccggg ggcaccncgg 600

134 aattttggga acccgggggg ggccttttt 629

136 <210> SEQ ID NO: 5

137 <211> LENGTH: 386

138 <212> TYPE: PRT

139 <213> ORGANISM: homo sapiens

141 <400> SEQUENCE: 5

142 Leu Phe Leu Glu Asn Leu Ile Lys Tyr Gly Ile Leu Val Asp Pro Ile

143 1 5 10 15

144 Gln Val Val Ser Leu Phe Leu Lys Asp Pro Tyr Ser Trp Pro Ala Pro

145 20 25 30

146 Cys Leu Val Ile Ala Ala Asn Val Phe Ala Val Ala Ala Phe Gln Val

147 35 40 45

148 Glu Lys Arg Leu Ala Val Gly Ala Leu Thr Glu Gln Ala Gly Leu Leu

149 50 55 60

150 Leu His Val Ala Asn Leu Ala Thr Ile Leu Cys Phe Pro Ala Ala Val

151 65 70 75 80

152 Val Leu Leu Val Glu Ser Ile Thr Pro Val Gly Ser Leu Leu Ala Leu

153 85 90 95

154 Met Ala His Thr Ile Leu Phe Leu Lys Leu Phe Ser Tyr Arg Asp Val

155 100 105 110

156 Asn Ser Trp Cys Arg Arg Ala Arg Ala Lys Ala Ala Ser Ala Gly Lys

157 115 120 125

158 Lys Ala Ser Ser Val Ala Ala Pro His Thr Val Ser Tyr Pro Asp Asn

159 130 135 140

160 Leu Thr Tyr Arg Asp Leu Tyr Tyr Phe Leu Phe Ala Pro Thr Leu Cys

161 145 150 155 160

162 Tyr Glu Leu Asn Phe Pro Arg Ser Pro Arg Ile Arg Lys Arg Phe Leu

163 165 170 175

164 Leu Arg Arg Ile Leu Glu Met Leu Phe Thr Gln Leu Gln Val Gly

165 180 185 190

166 Leu Ile Gln Gln Trp Met Val Pro Thr Ile Gln Asn Ser Met Lys Pro

RAW SEQUENCE LISTING

DATE: 07/24/2002

PATENT APPLICATION: US/10/040,315A

TIME: 11:25:15

Input Set : D:\seqlist.txt

Output Set: N:\CRF3\07242002\J040315A.raw

```

167          195          200          205
168 Phe Lys Asp Met Asp Tyr Ser Arg Ile Ile Glu Arg Leu Leu Lys Leu
169      210          215          220
170 Ala Val Pro Asn His Leu Ile Trp Leu Ile Phe Phe Tyr Trp Leu Phe
171 225      230          235          240
172 His Ser Cys Leu Asn Ala Val Ala Glu Leu Met Gln Phe Gly Asp Arg
173      245          250          255
174 Glu Phe Tyr Arg Asp Trp Trp Asn Ser Glu Ser Val Thr Tyr Phe Trp
175      260          265          270
176 Gln Asn Trp Asn Ile Pro Val His Lys Trp Cys Ile Arg His Phe Tyr
177      275          280          285
178 Lys Pro Met Leu Arg Arg Gly Ser Ser Lys Trp Met Ala Arg Thr Gly
179      290          295          300
180 Val Phe Leu Ala Ser Ala Phe Phe His Glu Tyr Leu Val Ser Val Pro
181 305      310          315          320
182 Leu Arg Met Phe Arg Leu Trp Ala Phe Thr Gly Met Met Ala Gln Ile
183      325          330          335
184 Pro Leu Ala Trp Phe Val Gly Arg Phe Phe Gln Gly Asn Tyr Gly Asn
185      340          345          350
186 Ala Ala Val Trp Leu Ser Leu Ile Ile Gly Gln Pro Ile Ala Val Leu
187      355          360          365
188 Met Tyr Val His Asp Tyr Tyr Val Leu Asn Tyr Glu Ala Pro Ala Ala
189      370          375          380
190 Glu Ala
191 385
193 <210> SEQ ID NO: 6
194 <211> LENGTH: 488
195 <212> TYPE: PRT
196 <213> ORGANISM: homo sapiens
198 <400> SEQUENCE: 6
199 Met Gly Asp Arg Gly Ser Ser Arg Arg Arg Arg Thr Gly Ser Arg Pro
200 1          5          10          15
201 Ser Ser His Gly Gly Gly Gly Pro Ala Ala Ala Glu Glu Glu Val Arg
202      20          25          30
203 Asp Ala Ala Ala Gly Pro Asp Val Gly Ala Ala Gly Asp Ala Pro Ala
204      35          40          45
205 Pro Ala Pro Asn Lys Asp Gly Asp Ala Gly Val Gly Ser Gly His Trp
206      50          55          60
207 Glu Leu Arg Cys His Arg Leu Gln Asp Ser Leu Phe Ser Ser Asp Ser
208      65          70          75          80
209 Gly Phe Ser Asn Tyr Arg Gly Ile Leu Asn Trp Cys Val Val Met Leu
210      85          90          95
211 Ile Leu Ser Asn Ala Arg Leu Phe Leu Glu Asn Leu Ile Lys Tyr Gly
212      100          105          110
213 Ile Leu Val Asp Pro Ile Gln Val Val Ser Leu Phe Leu Lys Asp Pro
214      115          120          125
215 His Ser Trp Pro Ala Pro Cys Leu Val Ile Ala Ala Asn Val Phe Ala
216      130          135          140
217 Val Ala Ala Phe Gln Val Glu Lys Arg Leu Ala Val Gly Ala Leu Thr

```

RAW SEQUENCE LISTING

DATE: 07/24/2002

PATENT APPLICATION: US/10/040,315A

TIME: 11:25:15

Input Set : D:\seqlist.txt

Output Set: N:\CRF3\07242002\J040315A.raw

```

218 145 150 155 160
219 Glu Gln Ala Gly Leu Leu Leu His Val Ala Asn Leu Ala Thr Ile Leu
220 165 170 175
221 Cys Phe Pro Ala Ala Val Val Leu Leu Val Glu Ser Ile Thr Pro Val
222 180 185 190
223 Gly Ser Leu Leu Ala Leu Met Ala His Thr Ile Leu Phe Leu Lys Leu
224 195 200 205
225 Phe Ser Tyr Arg Asp Val Asn Ser Trp Cys Arg Arg Ala Arg Ala Lys
226 210 215 220
227 Ala Ala Ser Ala Gly Lys Lys Ala Ser Ser Ala Ala Ala Pro His Thr
228 225 230 235 240
229 Val Ser Tyr Pro Asp Asn Leu Thr Tyr Arg Asp Leu Tyr Tyr Phe Leu
230 245 250 255
231 Phe Ala Pro Thr Leu Cys Tyr Glu Leu Asn Phe Pro Arg Ser Pro Arg
232 260 265 270
233 Ile Arg Lys Arg Phe Leu Leu Arg Arg Ile Leu Glu Met Leu Phe Phe
234 275 280 285
235 Thr Gln Leu Gln Val Gly Leu Ile Gln Gln Trp Met Val Pro Thr Ile
236 290 295 300
237 Gln Asn Ser Met Lys Pro Phe Lys Asp Met Asp Tyr Ser Arg Ile Ile
238 305 310 315 320
239 Glu Arg Leu Leu Lys Leu Ala Val Pro Asn His Leu Ile Trp Leu Ile
240 325 330 335
241 Phe Phe Tyr Trp Leu Phe His Ser Cys Leu Asn Ala Val Ala Glu Leu
242 340 345 350
243 Met Gln Phe Gly Asp Arg Glu Phe Tyr Arg Asp Trp Trp Asn Ser Glu
244 355 360 365
245 Ser Val Thr Tyr Phe Trp Gln Asn Trp Asn Ile Pro Val His Lys Trp
246 370 375 380
247 Cys Ile Arg His Phe Tyr Lys Pro Met Leu Arg Arg Gly Ser Ser Lys
248 385 390 395 400
249 Trp Met Ala Arg Thr Gly Val Phe Leu Ala Ser Ala Phe Phe His Glu
250 405 410 415
251 Tyr Leu Val Ser Val Pro Leu Arg Met Phe Arg Leu Trp Ala Phe Thr
252 420 425 430
253 Gly Met Met Ala Gln Ile Pro Leu Ala Trp Phe Val Gly Arg Phe Phe
254 435 440 445
255 Gln Gly Asn Tyr Gly Asn Ala Ala Val Trp Leu Ser Leu Ile Ile Gly
256 450 455 460
257 Gln Pro Ile Ala Val Leu Met Tyr Val His Asp Tyr Tyr Val Leu Asn
258 465 470 475 480
259 Tyr Glu Ala Pro Ala Ala Glu Ala
260 485
262 <210> SEQ ID NO: 7
263 <211> LENGTH: 498
264 <212> TYPE: PRT
265 <213> ORGANISM: mus musculus
267 <400> SEQUENCE: 7
268 Met Gly Asp Arg Gly Gly Ala Gly Ser Ser Arg Arg Arg Arg Thr Gly

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 07/24/2002
PATENT APPLICATION: US/10/040,315A TIME: 11:25:16

Input Set : D:\seqlist.txt
Output Set: N:\CRF3\07242002\J040315A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:4; N Pos. 455,464,467,475,497,500,508,514,519,536,543,544,576,583,584

Seq#:4; N Pos. 597

VERIFICATION SUMMARY

DATE: 07/24/2002

PATENT APPLICATION: US/10/040,315A

TIME: 11:25:16

Input Set : D:\seqlist.txt

Output Set: N:\CRF3\07242002\J040315A.raw

L:131 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:420
L:132 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:480
L:133 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:540